3.0 Quality Control

As is the case with most real-time meteorological/ocean observations, the real-time QC of wind observations can be extremely challenging. Events such as fast moving fronts, microbursts, and tropical cyclones must be considered when determining acceptable data thresholds. Human involvement is therefore important to ensure that solid scientific principles are applied to data evaluation so that good data are not discarded and bad data are not distributed (e.g., selection of appropriate thresholds and examination of data flagged as questionable).

To conduct real-time QC on wind observations, the first pre-requisite is to understand the science and context within which the measurements are being conducted. For example and as was discussed in section 2.2.3, sensors can be deployed in a number of ways. Each deployment method imposes the need for specific QC methods. Real-time wind data should have these main attributes: accurate time, speed, direction, and gust measurements.

This manual focuses specifically on the QC of real-time data, but there are limitations. For example, gradual calibration changes or slow system response variations (sensor drift) cannot be detected or corrected in real time. Therefore, delayed-mode approaches are done through comparison with collocated observations (e.g., satellite data). Drift correction to wind measurements during post-processing is highly unlikely to occur even if a valid post-recovery calibration could be obtained. Drift is often caused by worn bearings and corrosion, potentiometer pad degradation, and to a lesser extent, the aging of electronic components—e.g., those compensating for temperature dependencies, which are now less of a problem thanks to advances in digital circuitry. These gradual changes affect different systems in different ways (e.g., an impellor has a higher starting speed caused by corroded bearings). Another limitation is the ability of some data providers to backfill data gaps. In both of these examples, the observations are not considered to be real time for purposes of QC checks. (However, in some sophisticated 24/7 QC operations, real-time dissemination may be switched from one sensor to another based on real-time QC flags.)

Observations are time ordered, and the most recent observation is n_0, preceded by a value at n_{-1}, and so on backwards in time. The focus of the real-time QC is primarily on observations n_0, n_{-1}, and n_{-2}.

3.1 QC Flags

Data are evaluated using QC tests, and the results of those tests are recorded by inserting flags in the data files. Table 3-1 provides the set of flags and associated descriptions adopted by the International Oceanographic Data and Information Exchange in 2013 and subsequently by U.S. IOOS. Additional flags may be incorporated to provide more detailed information to assist with troubleshooting. For example, an observation may fail the wind speed neighbor test and be flagged as having failed. If the data failed the wind speed neighbor test because the observation is too low, a second-tier "failed low" flag may indicate that the value was lower than allowed by a preset threshold. Such detailed flags primarily support maintenance efforts and are presently beyond U.S. IOOS requirements for QC of real-time data. However, all flags should be identified and defined in the data's metadata.

Further post-processing of the data may yield different conclusions from those reached during initial assessments. Flags set in real time should not be changed, ensuring that historical documentation is preserved. Results from post processing should generate another set of flags corresponding to a revised version of the data.

Table 3-1. Flags for real-time data (UNESCO 2013)

Flag	Description
Pass=1	Data have passed critical real-time quality control tests and are deemed adequate for use as preliminary data.
Not Evaluated=2	Data have not been QC-tested, or the information on quality is not available.
Suspect or Of High Interest=3	Data are considered to be either suspect or of high interest to data providers and users. They are flagged suspect to draw further attention to them by operators.
Fail=4	Data are considered to have failed one or more critical real-time QC checks. If they are disseminated at all, it should be readily apparent that they are not of acceptable quality.
Missing Data=9	Data are missing; used as a placeholder.

3.2 Test Hierarchy

This section outlines eleven real-time QC tests that are required, recommended, or suggested for wind measurements. Operators should also consider that some of these tests can be carried out within the instrument, where thresholds can be defined in configuration files. Although more tests may imply a more robust QC effort, there are many reasons operators could use to justify not conducting some tests. In those cases, operators need only to document reasons these tests do not apply to their observations. Such flexibility is needed to support the U.S. IOOS effort, since the number of tests conducted and the justification for not applying some tests are useful for evaluating an operator's skill levels. Tests are listed in table 3-2 and are divided into three groups: those that are required, strongly recommended, or suggested. However, for some critical real-time applications with high risk operations, it may be advisable to invoke all groups.

Table 3-2. QC Tests in order of implementation and hierarchy

Group 1 *Required*	Test 1 Test 2 Test 3 Test 4 Test 5	Timing/Gap Test Syntax Test Location Test Gross Range Test Climatology Test
Group 2 *Strongly Recommended*	Test 6 Test 7 Test 8	Spike Test Rate of Change Test Flat Line Test
Group 3 *Suggested*	Test 9 Test 10 Test 11	Multi-Variate Test Attenuated Signal Test Neighbor Test

3.3 QC Test Descriptions

A variety of tests can be performed on the sensor measurements to evaluate data quality. Testing the timely arrival and integrity of the data transmission itself is a first step. If the data are corrupted during transmission, further testing may be irrelevant. The checks defined in these eleven tests evaluate data through various comparisons to other data and to the expected conditions in the given environment. The tests listed in this section presume a time-ordered series of observations and denote the most recent observation as previously described.

Some effort will be needed to select the best thresholds, which are determined at the operator level and may require multiple iterations of trial and error before final selections are made. A successful QC effort is highly dependent upon selection of the proper thresholds, which should not be determined arbitrarily but can be based on historical knowledge or statistics derived from recently acquired data. Although this manual provides some guidance for selecting thresholds based on input from various operators, it is assumed that operators have the necessary expertise and interest in selecting the proper thresholds to maximize the value of their QC effort. Operators should openly provide thresholds as metadata for user support. The selection of wind thresholds may be dependent upon the real-time application, (e.g., onset of a coastal sea breeze or observation of a hurricane maximum gust). This shared information will help U.S. IOOS to document standardized thresholds that will be included in future releases of this manual.

3.3.1 Applications of QC Tests to Wind Sensors

These eleven tests require operators to select a variety of thresholds. Examples are provided in the following test tables; however, operators are in the best position to determine the appropriate thresholds for their operations. Wind speed (WS) is used in the descriptions and examples, but the tests apply equally to direction and gust in most cases. A discontinuity in wind direction is caused when the wind veers through north, stepping from 359° to 0° and complicating the application of some of these tests. Operators may choose to conduct wind direction tests on the *u* and *v* wind direction components to circumvent the problem. Some tests rely on multiple data points most recently received to determine the quality of the latest data point. When this series of data points reveals that the entire group fails, the most recent data point is flagged, but the previous flags are not changed. This action supports the view that historical flags are generally not altered. The first example is in Test 8, the Flat Line Test, where this scenario will become clearer. The exception to the rule occurs for Test 6 Spike Check, where the most recent point must be flagged as “2 Not Evaluated” until the next point arrives and the spike check can be performed. For additional information regarding flags, see the *Manual for the Use of Real-Time Oceanographic Data Quality Control Flags* (U.S. IOOS 2014) posted on the U.S. IOOS QARTOD website.

Test 1 - Timing/Gap Test (Required)

Check for arrival of data.

Test determines that the most recent data point has been measured and received within the expected time window (TIM_INC) and has the correct time stamp (TIM_STMP).

Note: For those systems that do not update at regular intervals (Argos telemetry, for example), a large value for TIM_STMP can be assigned. The gap check is not a solution for all timing errors. Data could be measured or received earlier than expected. This test does not address all clock drift/jump issues.

Flags	Condition	Codable Instructions
Missing Data=9	Data have not arrived as expected.	If NOW – TIM_STMP > TIM_INC, flag = 9
Suspect=3	N/A	N/A
Pass=1	Applies for test pass condition.	N/A

Test Exception: None.

Test specifications to be established locally by the operator.
Example: TIM_INC= 1 hour

Test 2 - Syntax Test (Required)

Check to ensure that the message is structured properly.

Received data message (full message) contains the proper structure without any indicators of flawed transmission such as parity errors. Possible tests are: a) the expected number of characters (NCHAR) for fixed-length messages equals the number of characters received (REC_CHAR), or b) passes a standard parity bit check, cyclic redundancy check, etc. Many such syntax tests exist, and the user should select the best criteria for one or more syntax tests.

Capabilities for dealing with flawed messages vary among operators; some may have the ability to parse messages to extract data within the flawed message sentence before the flaw. A syntax check is performed only at the message level and not within the message content. In cases where a data record requires multiple messages, this check can be performed at the message level but is not used to check message content.

Flags	Condition	Codable Instructions
Fail=4	Data sentence cannot be parsed to provide a valid observation.	If REC_CHAR ≠ NCHAR, flag = 4
Suspect =3	N/A	N/A
Pass=1	Expected data sentence received; absence of parity errors.	N/A

Test Exception: None.

Test specifications to be established locally by the operator.
Example: NCHAR = 128

Test 3 - Location Test (Required)

Check for reasonable geographic location.		
Test checks that the reported present physical location (latitude/longitude) is within operator-determined limits. The location test(s) can vary from: 1) a simple invalid location, to 2) a more complex check for displacement (DISP) exceeding a distance limit RANGEMAX based upon a previous location and platform speed. Operators may also check for 3) erroneous locations based upon other criteria, such as reported positions over land, as appropriate.		
Flags	**Condition**	**Codable Instructions**
Fail=4	Invalid location	If |LAT| > 90 or |LONG| > 180, flag = 4
Suspect=3	Unlikely platform displacement	If DISP > RANGEMAX, flag = 3
Pass=1	Applies for test pass condition.	N/A
Test Exception: Test does not apply to fixed deployments when no location is transmitted.		
Test specifications to be established locally by the operator. **Example 1:** Impossible location, LAT or LONG exceeds mathematical limits. **Example 2:** Displacement DISP calculated between sequential position reports, RANGEMAX = 20 km. **Example 3:** Buoy position resides within land mask.		

Test 4 - Gross Range Test (Required)

Data point exceeds sensor or operator-selected min/max.		
All sensors have a limited output range, and this can form the most rudimentary gross range check. No values less than a minimum value or greater than the maximum value the sensor can output (SENSOR_MIN, SENSOR_MAX) are acceptable. Additionally, the operator can select a smaller span (USER_MIN, USER_MAX) based upon local knowledge or a desire to draw attention to extreme values. An obvious gross range check is wind direction 0-360°. **NOTE:** Operators may choose to flag as suspect values that exceed the calibration span but not the hardware limits (e.g., a value that sensor is not capable of producing).		
Flags	**Condition**	**Codable Instructions**
Fail=4	Reported value is outside of sensor span.	If WS_n < SENSOR_MIN, or WS_n > SENSOR_MAX, flag = 4
Suspect=3	Reported value is outside of user-selected span.	If WS_n < USER_MIN, or WS_n > USER_MAX, flag = 3
Pass=1	Applies for test pass condition	
Test Exception: None.		
Test specifications to be established locally by the operator. **Examples:** SENSOR_MAX = 100 m/s (limited by the manufacturer firmware, for example) SENSOR_MIN = 0 m/s USER_MAX = 75 m/s USER_MIN = - 0 m/s		

Test 5 - Climatology Test (Required)

Test that data point falls within seasonal expectations.

This test is a variation on the gross range check, where the gross range Season_MAX and Season_MIN are adjusted monthly, seasonally, or at some other operator-selected time period (TIM_TST). Expertise of the local operator using long historical records is the best method to determine reasonable seasonal averages - longer time series permit more refined identification of appropriate thresholds. Additional climatology guidance is available at http://www.ncdc.noaa.gov/societal-impacts/wind/mean/2014/4, http://numbat.coas.oregonstate.edu/cogow, http://iridl.ldeo.columbia.edu/maproom/Global/Climatologies/Vector_Winds.html, and from the NCEP/NCAR Reanalysis 1, 2 and 3 (now CFSR).

Flags	Condition	Codable Instructions
Fail=4	Because of the potential for extreme wind speeds, gusts, and directional variability without regard to season, no fail flag is identified for this test.	N/A
Suspect=3	Reported value is outside the operator-identified climatology window.	If WS_n < Season_MIN or WS_n > Season_MAX, flag = 3
Pass=1	Applies for test pass condition.	N/A

Test Exception: None.

Test specifications to be established locally by operator: A seasonal matrix of WS_{max} and WS_{min} values at all TIM_TST intervals.
Examples: SPRING_MIN = 0 m/s, SPRING_MAX = 60 m/s

Test 6 - Spike Test (Strongly Recommended)

<table>
<tr><td colspan="3">Data point n-1 exceeds a selected threshold relative to adjacent data points.</td></tr>
<tr><td colspan="3">This check is for single-value spikes, specifically the value at point n-1. Spikes consisting of more than one data point are difficult to capture, but their onset may be flagged by the rate of change test. The spike test consists of two operator-selected thresholds, THRSHLD_LOW and THRSHLD_HIGH. Adjacent data points (n_{-2} and n_0) are averaged to form a spike reference (SPK_REF). The absolute value of the spike is tested to capture positive and negative spikes. Large spikes are easier to identify as outliers and flag as failures. Smaller spikes may be real and are only flagged suspect. The thresholds may be fixed values or dynamically established (for example, a multiple of the standard deviation over an operator-selected period).

An alternative is a third difference test defined as $Diff_n = WS_{n-3} - 3^* WS_{n-2} + 3^* WS_{n-1} - WS_n$.</td></tr>
<tr><td>Flags</td><td>Condition</td><td>Codable Instructions</td></tr>
<tr><td>Fail=4</td><td>High spike threshold exceeded.</td><td>If $|WS_{n-1}$ - SPK_REF$|$ > THRSHLD_HIGH, flag = 4</td></tr>
<tr><td>Suspect=3</td><td>Low spike threshold exceeded.</td><td>If $|WS_{n-1}$ - SPK_REF$|$ > THRSHLD_LOW and
$|WS_{n-1}$ - SPK_REF$| \leq$ THRSHLD_HIGH, flag = 3</td></tr>
<tr><td>Pass=1</td><td>Applies for test pass condition.</td><td>N/A</td></tr>
<tr><td colspan="3">Test Exception: None.</td></tr>
<tr><td colspan="3">Test specifications to be established locally by the operator.
Examples: THRSHLD_LOW = 20 m/s, THRSHLD_HIGH = 40 m/s</td></tr>
</table>

Note: For one-minute sampling, a threshold proportional to the 97th or 98th percentile of first differences is effective given enough recent data to robustly calculate this threshold. This flexible standard is particularly useful for ships, which can traverse a wide range of conditions and sensors in areas with large synoptic or seasonal scale variability.

Test 7 - Rate of Change Test (Strongly Recommended)

Excessive rise/fall test.

This test inspects the time series for a time rate of change that exceeds a threshold value identified by the operator. Wind speed, direction, and gust values can change substantially over short periods in all locations, hindering the value of this test. A balance must be found between a threshold set too low, which triggers too many false alarms, and one set too high, making the test ineffective. Test implementation can be challenging. Upon failure, it is unknown which point is bad. Further, upon failing a data point, it remains to be determined how the next iteration can be handled. The following suggests one approach to implementation of a threshold:

> The rate of change between WS_{n-1} and WS_n must be less than three standard deviations (3*SD) of first differences. The local operator can determine both the number of SDs (N_DEV) and the period over which the SD is calculated (TIM_DEV).

Flags	Condition	Codable Instructions
Fail=4	No fail flag is identified for this test.	N/A
Suspect=3	The rate of change exceeds the selected threshold.	If $\lvert WS_n - WS_{n-1} \rvert$ > N_DEV*SD, flag = 3
Pass=1	Applies for test pass condition.	N/A

Test Exception: None.

Test specifications to be established locally by operator.
Examples: N_DEV = 3, TIM_DEV = 8 hours.

Test 8 - Flat Line Test (Strongly Recommended)

Invariant value.		
When some sensors and/or data DCPs fail, the result can be a continuously repeated observation of the same value. This test compares the present observation *n* to a number (REP_CNT_FAIL or REP_CNT_SUSPECT) of previous observations. Observation *n* is flagged if it has the same value as previous observations within a tolerance value, EPS, to allow for numerical round-off error. Note that historical flags are not changed.		
Flags	**Condition**	**Codable Instructions**
Fail=4	When the five most recent observations are equal, WS_n is flagged fail.	Flag=4 For *i*=1,REP_CNT_FAIL If $\|WS_n - WS_{n-i}\|$ > EPS then flag = 1 end if
Suspect=3	It is possible but unlikely that the present observation and the two previous observations would be equal. When the three most recent observations are equal, WS_n is flagged suspect.	If flag=4, end Flag=3 For *i*=1,REP_CNT_SUSPECT If $\|WS_n - WS_{n-i}\|$> EPS then flag = 1 end if
Pass=1	Applies for test pass condition.	N/A
Test Exception: None.		
Test specifications to be established locally by the operator. **Examples:** REP_CNT_FAIL = 5, REP_CNT_SUSPECT= 3, EPS = 0.5 m/s, in some instances, EPS=0 might apply.		

Test 9 - Multi-Variate Test (Suggested)

This is an advanced family of tests, starting with the simpler test described here and anticipating growth towards full co-variance testing in the future. It is doubtful that anyone is conducting tests such as these in real time. As these tests are developed and implemented, they should be documented and standardized in later versions of this manual.

Comparison to other variables.		
This example pairs rate of change tests as described in Test 7. The WS rate of change test is conducted with a more restrictive threshold (N_WSMV_DEV). If this test fails, a second rate of change test operating on a second variable (barometric pressure [BP], for example) is conducted. The absolute value rate of change should be tested, since the relationship between WS and the second variable may be indeterminate. If the rate of change test on the second variable fails to exceed a threshold (e.g., an anomalous step is found in WS and is lacking in barometric pressure), then the WS_n value is flagged.		
Flags	**Condition**	**Codable Instructions**
Fail=4	No fail flag is identified for this test.	N/A
Suspect=3	WS_n fails the rate of change and the second variable (barometric pressure, for example) does not exceed the rate of change.	If $\lvert WS_n – WS_{n-1} \rvert$ > N_WSMV_DEV*SD_WS AND $\lvert BP_n – BP_{n-1} \rvert$ < N_BP_DEV*SD_BP, flag = 3
Pass=1		N/A
Test Exception: None.		
Test specifications to be established locally by the operator. **Examples:** N_WSMV_DEV = 2, N_BP_DEV=2, TIM_DEV = 8 hours		

NOTE: In a more complex case, more than one secondary rate of change test can be conducted. Wind direction or air temperature could be possible secondary candidates to be checked for anomalous rate of change values. In this case, a knowledgeable operator may elect to assign a pass flag to a high rate of change observation when any one of the secondary variables also exhibits a high rate of change. Such tests border on modeling, should be carefully considered, and may be beyond the scope of this effort.

The QARTOD wind committee recognized the high value in full co-variance testing but also noted the challenges. Such testing remains to be a research project not yet ready for operational implementation.

Test 10 - Attenuated Signal Test (Suggested)

A test for inadequate variation of the time series.

A common sensor failure mode can provide a data series that is nearly but not exactly a flat line. Badly worn bearings, a failed grounding wire, signal crosstalk, or inadequate wire shielding might cause such a failure. This test inspects for an SD value or a range variation (MAX-MIN) value that fails to exceed threshold values (MIN_VAR_WARN, MIN_VAR_FAIL) over a selected time period (TST_TIM).

Flags	Condition	Codable Instructions
Fail=4	Variation fails to meet the minimum threshold MIN_VAR_FAIL.	If During TST_TIM, SD <MIN_VAR_FAIL, or During TST_TIM, MAX-MIN <MIN_VAR_FAIL, flag = 4
Suspect=3	Variation fails to meet the minimum threshold MIN_VAR_WARN.	If During TST_TIM, SD <MIN_VAR_WARN, or During TST_TIM, MAX-MIN <MIN_VAR_WARN, flag = 3
Pass=1	Applies for test pass condition.	N/A

Test Exception: None.

Test specifications to be established locally by the operator.
Examples: TST_TIM = 12 hours
MIN_VAR_WARN= ?, MIN_VAR_FAIL= ?

Note: This type of failure mode is rare for most anemometers, occurring at very low speeds for cup anemometers. A related problem occurs with sonic anemometers when droplets of water bead on the transmitter and receiver. The speed of sound is faster in water, complicating the interpretation of the observations. The consequences of this problem are not easily identified in wind speeds but can be a serious problem if the instrument is used to measure a momentum flux.

Test 11 - Neighbor Test (Suggested)

Comparison to nearby sensors.

This check has the potential to be the most useful test when a nearby second sensor is determined to have a similar response.

Ideally, redundant sensors utilizing different technology would be co-located and alternately serviced at different intervals. This close neighbor would provide the ultimate QC check, but cost prohibits such a deployment in most cases.

However, there are few instances where a second sensor is sufficiently proximate to provide a useful QC check. Wind observations are more readily compared to adjacent sites than many non-conservative observations (such as dissolved oxygen, for example), and this test should not be overlooked where it may have application.

This test is the same as Test 9), *Multi-Variate Check – comparison to other variables* where the second variable is the second sensor. The selected thresholds depend entirely upon the relationship between the two sensors as determined by the local knowledge of the operator.

In the instructions and examples below, data from one site (WS1) are compared to a second site (WS2). The standard deviation for each site (SD1, SD2) is calculated over the period (TIM_DEV) and multiplied as appropriate (N_WS1_DEV for site WS1) to calculate the rate of change threshold. Note that an operator could also choose to use the same threshold for each site, since the sites are presumed to be similar. A unique and highly valuable version of the neighbor check is the surrogate use of wind forecasts. These 'virtual neighbor' constructs offer a QC check that is also presumed to be similar—again, within user-selected thresholds.

Flags	**Condition**	**Codable Instructions**
Fail=4	No fail flag is identified for this test.	N/A
Suspect=3	$WS1_n$ fails the rate of change and the second sensor $WS2_n$ does not exceed the rate of change.	If $\lvert WS1_n – WS1_{n-1} \rvert$ > N_WS1_DEV*SD1 AND $\lvert WS2_n – WS2_{n-1} \rvert$ < N_WS2_DEV*SD2, flag = 3
Pass=1		N/A

Test Exception: There is no adequate neighbor.

Test specifications to be established locally by the operator.
Examples: N_WS1_DEV = 2, N_WS2_DEV=2, TIM_DEV = 8 hours